Clustering methods- Results

Saritha Kodikara

February 11, 2022

Contents

1 Simulation Results- Original data

In here, we present the clustering accuracy for four clustering method ("PCA", "DTW", "Agglomerative", "Kmedoids") with original linear mixed model (LMM) profiles.

```
temp <-list.files(path="Data/", pattern="*.Rdata")

for(i in 1:3){
  clu_output<-data.frame()
  modelledData<-readRDS(paste0("Data/",temp[i]))

  for(j in 1:length(modelledData)){
    lmmData<- modelledData[[j]]
    methods_accuracy<-cluster_Methods_acc(lmmData)
    clu_output<-rbind(clu_output,methods_accuracy)
    names(clu_output)<-c("PCA", "DTW",
                         "Agglomerative", "Kmedoids")
  }
  write.csv(clu_output, file =  paste0("Results/Org_",i,".csv"))
}
```

```
temp <-list.files(path="Results/",pattern= "Org_")
for(i in 1:3){
  clRes<-read.csv(paste0("Results/",temp[i]))[,-1]
  nam <- paste("plot", i, sep = "_")
  assign(nam, clusBoxplot(clRes, plotTitle = ""))
}

P<-plot_1+plot_2+plot_3+ plot_layout(guides = "collect")


temp <-list.files(path="Data/", pattern="*.Rdata")
for(i in 1:3){
```

```
  clData<-readRDS(paste0("Data/",temp[i]))[[2]]
  nam <- paste("plot", i, sep = "_")
  if(i==1) n="Noise=0.5"
  if(i==2) n="Noise=1.5"
  if(i==3) n="Noise=3"
  assign(nam, plotFunction(clData, plotTitle = n))
}

Q<-plot_1+plot_2+plot_3+ plot_layout(guides = "collect")

Q/P
```

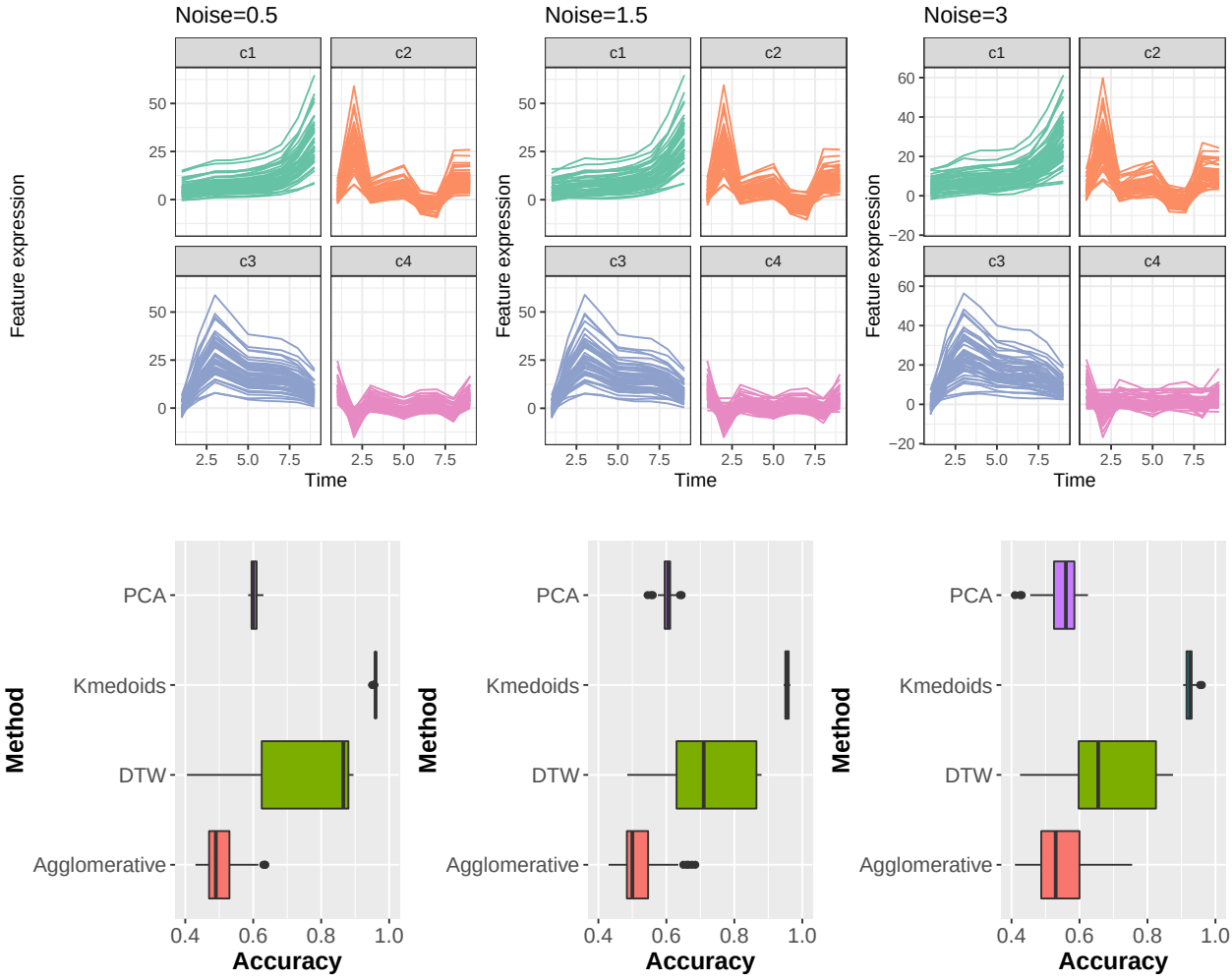

Figure 1: Figure 1, Clustering accuracy for PCA, K-medoid, DTW and Agglomerative clustering for original LMMS data.

2 Simulation Results- Centered data

In here, we present the clustering accuracy for four clustering method ("PCA", "DTW", "Agglomerative", "Kmedoids") with centered LMM profiles.

```
temp <-list.files(path="Data/", pattern="*.Rdata")

for(i in 1:3){
  clu_output<-data.frame()
  modelledData<-readRDS(paste0("Data/",temp[i]))

  for(j in 1:length(modelledData)){
    Centered.lmmData<-scale(modelledData[[j]],scale=FALSE)
    methods_accuracy<-cluster_Methods_acc(Centered.lmmData)
    clu_output<-rbind(clu_output,methods_accuracy)
    names(clu_output)<-c("PCA", "DTW",
                         "Agglomerative", "Kmedoids")
  }
  write.csv(clu_output, file =  paste0("Results/ Center_",i,".csv"))
}
```

```
temp <-list.files(path="Results/",pattern="Center_")
for(i in 1:3){
  clRes<-read.csv(paste0("Results/",temp[i]))[,-1]
  nam <- paste("plot", i, sep = "_")
  assign(nam, clusBoxplot(clRes, plotTitle = ""))
}

P<-plot_1+plot_2+plot_3+ plot_layout(guides = "collect")


temp <-list.files(path="Data/", pattern="*.Rdata")
for(i in 1:3){
  clData<-readRDS(paste0("Data/",temp[i]))[[2]]
  nam <- paste("plot", i, sep = "_")
  if(i==1) n="Noise=0.5"
  if(i==2) n="Noise=1.5"
  if(i==3) n="Noise=3"
  assign(nam, plotFunction(clData, plotTitle = n,center = TRUE))
}

Q<-plot_1+plot_2+plot_3+ plot_layout(guides = "collect")

Q/P
```

3 Simulation Results- Scaled data

In here, we present the clustering accuracy for four clustering method (“PCA”, “DTW”, “Agglomerative”, “Kmedoids”) with scaled LMM profiles.

```
temp <-list.files(path="Data/", pattern="*.Rdata")

for(i in 1:3){
  clu_output<-data.frame()
  modelledData<-readRDS(paste0("Data/",temp[i]))

  for(j in 1:length(modelledData)){
    Scaled.lmmData<-scale(modelledData[[j]],center =FALSE)
```

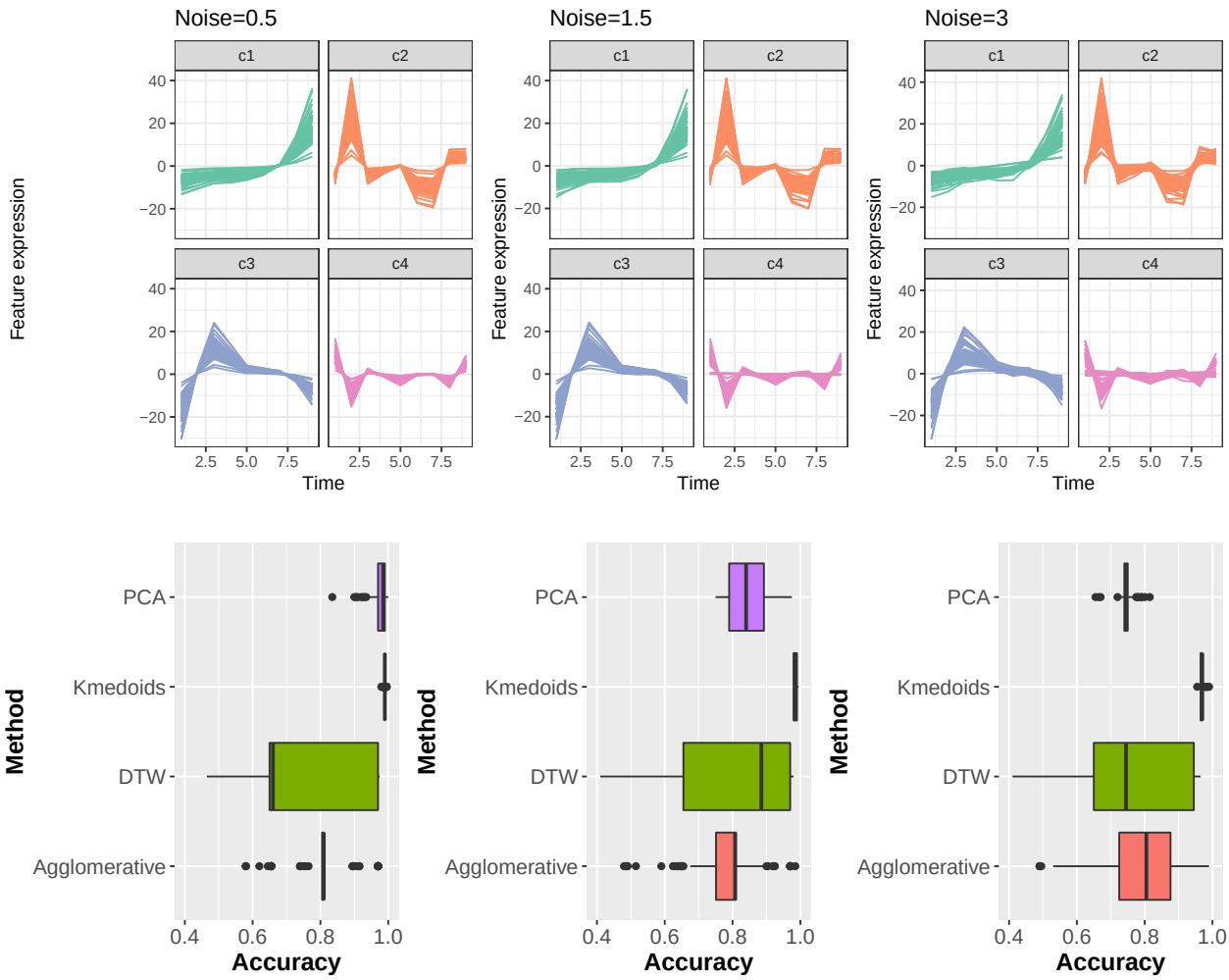

Figure 2: Figure 2, Clustering accuracy for PCA, K-medoid, DTW and Agglomerative clustering for centered LMMS data.

```
    methods_accuracy<-cluster_Methods_acc(Scaled.lmmData)
    clu_output<-rbind(clu_output,methods_accuracy)
    names(clu_output)<-c("PCA", "DTW",
                         "Agglomerative", "Kmedoids")
  }
  write.csv(clu_output, file =  paste0("Results/ Scale_",i,".csv"))
}
```

```
temp <-list.files(path="Results/",pattern="Scale_")
for(i in 1:3){
  clRes<-read.csv(paste0("Results/",temp[i]))[,-1]
  nam <- paste("plot", i, sep = "_")
  assign(nam, clusBoxplot(clRes, plotTitle = ""))
}

P<-plot_1+plot_2+plot_3+ plot_layout(guides = "collect")


temp <-list.files(path="Data/", pattern="*.Rdata")
for(i in 1:3){
  clData<-readRDS(paste0("Data/",temp[i]))[[2]]
  nam <- paste("plot", i, sep = "_")
  if(i==1) n="Noise=0.5"
  if(i==2) n="Noise=1.5"
  if(i==3) n="Noise=3"
  assign(nam, plotFunction(clData, plotTitle = n, scale = TRUE))
}

Q<-plot_1+plot_2+plot_3+ plot_layout(guides = "collect")

Q/P
```

4 Simulation Results- Centered and Scaled data

In here, we present the clustering accuracy for four clustering method ("PCA", "DTW", "Agglomerative", "Kmedoids") with centered and scaled LMM profiles.

```
temp <-list.files(path="Data/", pattern="*.Rdata")

for(i in 1:3){
  clu_output<-data.frame()
  modelledData<-readRDS(paste0("Data/",temp[i]))

  for(j in 1:length(modelledData)){
    CenteredScaled.lmmData<-scale(modelledData[[j]])
    methods_accuracy<-cluster_Methods_acc(CenteredScaled.lmmData)
    clu_output<-rbind(clu_output,methods_accuracy)
    names(clu_output)<-c("PCA", "DTW",
                         "Agglomerative", "Kmedoids")
  }
  write.csv(clu_output, file =  paste0("Results/ Cen_Scal_",i,".csv"))
}
```

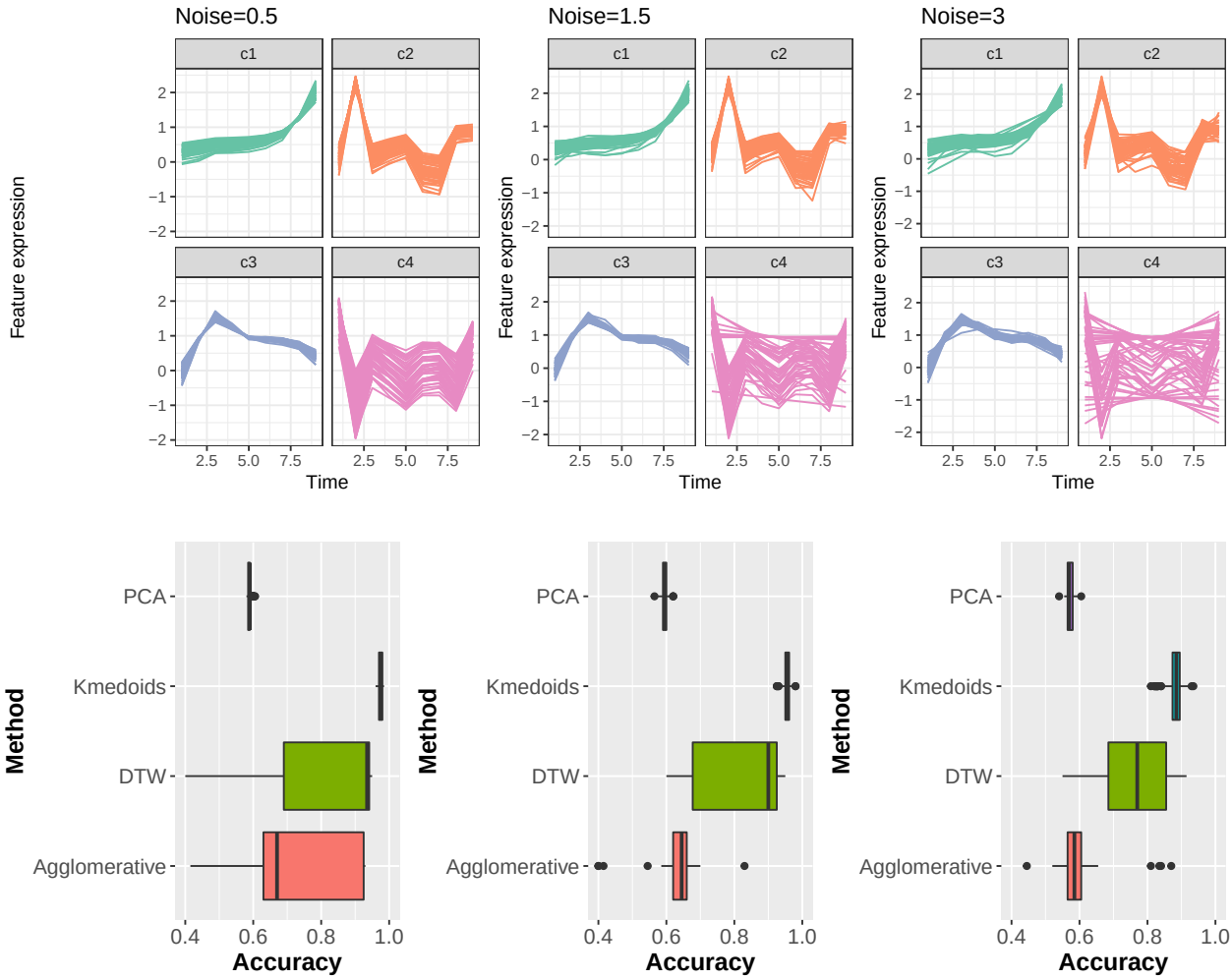

Figure 3: Figure 2, Clustering accuracy for PCA, K-medoid, DTW and Agglomerative clustering for scaled LMMS data.

```
temp <-list.files(path="Results/",pattern="Cen_Scal_")
for(i in 1:3){
  clRes<-read.csv(paste0("Results/",temp[i]))[,-1]
  nam <- paste("plot", i, sep = "_")
  assign(nam, clusBoxplot(clRes, plotTitle = ""))
}

P<-plot_1+plot_2+plot_3+ plot_layout(guides = "collect")


temp <-list.files(path="Data/", pattern="*.Rdata")
for(i in 1:3){
  clData<-readRDS(paste0("Data/",temp[i]))[[2]]
  nam <- paste("plot", i, sep = "_")
  if(i==1) n="Noise=0.5"
  if(i==2) n="Noise=1.5"
  if(i==3) n="Noise=3"
  assign(nam, plotFunction(clData, plotTitle = n, scale = TRUE, center = TRUE))
}

Q<-plot_1+plot_2+plot_3+ plot_layout(guides = "collect")

Q/P
```

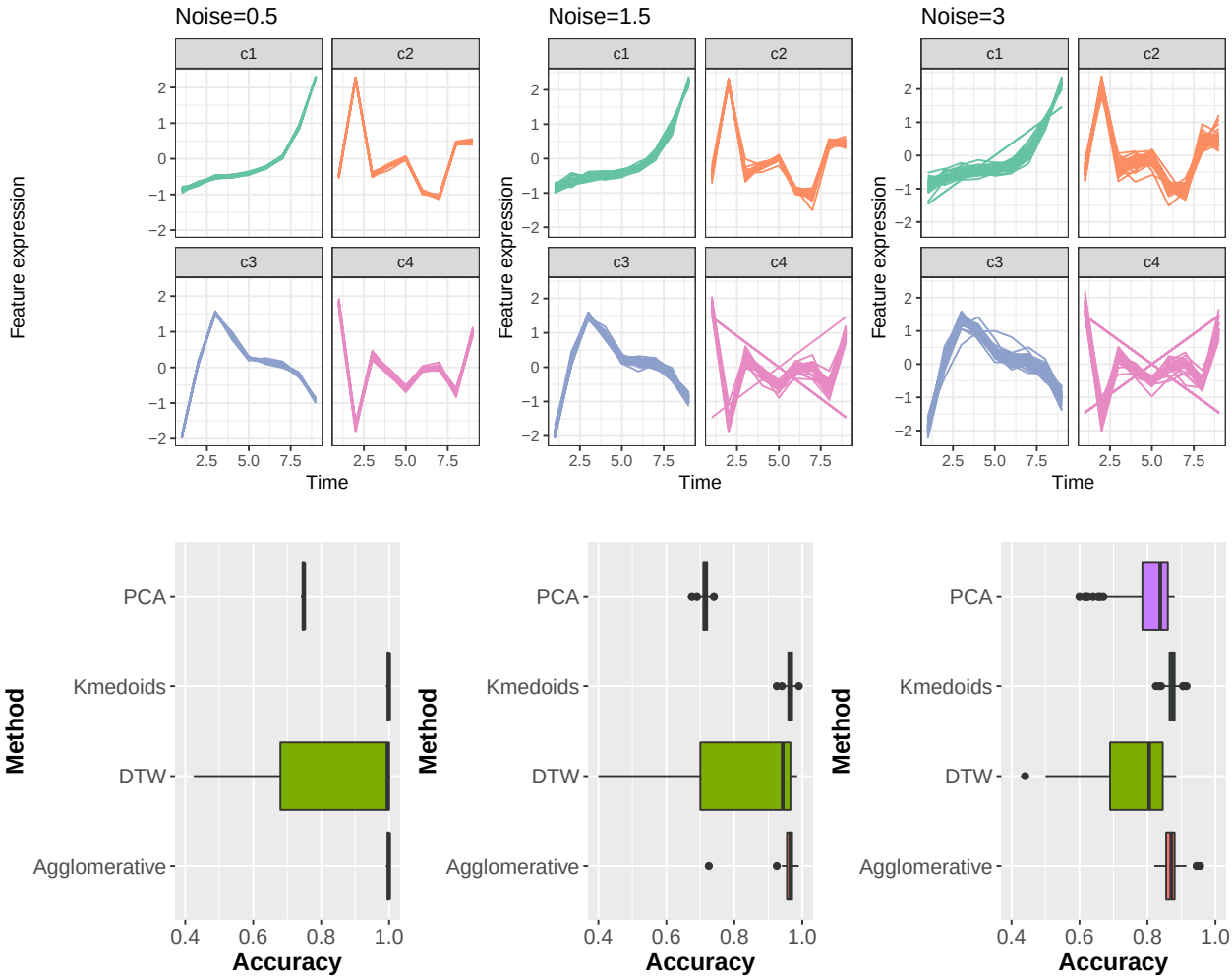

Figure 4: Figure 2, Clustering accuracy for PCA, K-medoid, DTW and Agglomerative clustering for centered and scaled LMMS data.